

Unlocking First Nations health information through data linkage

Walker, Jennifer; Pyper, E.; Jones, Carmen R.; Khan, Saba; Chong, N.; Legge, D.; Schull, M. J.; Henry, D.

Published in:
International Journal of Population Data Science

DOI:
[10.23889/ijpds.v3i1.450](https://doi.org/10.23889/ijpds.v3i1.450)

Licence:
CC BY-NC-ND

[Link to output in Bond University research repository.](#)

Recommended citation(APA):

Walker, J., Pyper, E., Jones, C. R., Khan, S., Chong, N., Legge, D., Schull, M. J., & Henry, D. (2018). Unlocking First Nations health information through data linkage. *International Journal of Population Data Science*, 3(1), [8]. <https://doi.org/10.23889/ijpds.v3i1.450>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

International Journal of Population Data Science

Journal Website: www.ijpds.org



Unlocking First Nations health information through data linkage

Walker, JD^{1,2*}, Pyper, E¹, Jones, CR³, Khan, S¹, Chong, N¹, Legge, D¹, Schull, MJ^{1,4,5,6}, and Henry, D^{1,7}

Submission History

| | |
|------------|------------|
| Submitted: | 21/12/2017 |
| Accepted: | 07/03/2018 |
| Published: | 22/05/2018 |

¹Institute for Clinical Evaluative Sciences, Toronto, ON, Canada

²School of Rural and Northern Health, Laurentian University, Sudbury, ON, Canada

³Chiefs of Ontario, Toronto, ON, Canada

⁴Evaluative Clinical Sciences, Trauma, Emergency & Critical Care Research Program, Sunnybrook Research Institute, Toronto, ON, Canada

⁵Department of Medicine, University of Toronto, ON, Canada

⁶Institute for Health Policy Management and Evaluation, University of Toronto, ON, Canada

⁷Centre for Research in Evidence-Based Practice, Bond University, Gold Coast, QLD, Australia

Abstract

Introduction

The importance of Indigenous data sovereignty and Indigenous-led research processes is increasingly being recognized in Canada and internationally. For First Nations in Ontario, Canada, access to routinely-collected demographic and health systems data is critical to planning and measuring health status and outcomes in their populations. Linkage of this data with the Indian Register (IR), under First Nations data governance, has unlocked data for use by First Nations organizations and communities.

Objectives

To describe the linkage of the IR database to the Ontario Registered Persons Database (RPDB) within the context of Indigenous data sovereignty principles.

Methods

Deterministic and probabilistic record linkage methods were used to link the IR to the RPDB. There is no established population of First Nations people living in Ontario with which we could establish a linkage rate. Accordingly, several approaches were taken to determine a denominator that would represent the total population of First Nations we would hope to link to the RPDB.

Results

Overall, 201,678 individuals in the national IR database matched to Ontario health records by way of the RPDB, of which 98,562 were female and 103,116 were male. Of those First Nations individuals linked to the RPDB, 90.2% (n=181,915) lived in Ontario when they first registered with IR, or were affiliated with an Ontario First Nation Community. The proportion of registered First Nations people linking to the RPDB improved across time, from 62.8% in the 1960s to 94.5% in 2012.

Conclusion

This linkage of the IR and RPDB has resulted in the creation of the largest First Nations health research study cohort in Canada. The linked data are being used by First Nations communities to answer questions that ultimately promote wellbeing, effective policy, and healing.

Introduction

Internationally, there is a strong movement towards Indigenous data sovereignty that highlights the importance of data governance and research processes that are Indigenous-led [1]. In Canada, First Nations established the principles of ownership, control, access and possession (OCAP[®]) of their data in the mid-1990s [2]. These principles embody collective rights to control the collection, use and storage of data about their populations and communities. However, the absence of mechanisms to meaningfully translate Indigenous data sovereignty principles into practice has led data holders to be highly restrictive in data access, resulting in data not being used to address First Nations research questions. For First Nations in Ontario, Canada's most populous province, access to routinely

collected demographic and health systems data is critical to planning and measuring progress as they undertake substantial efforts to improve the health status of their populations. A collaborative approach to data linkage is needed in order to fill these data gaps. The timing is critical as Canada is recognizing the enduring negative effects of colonial patterns of assimilation and exclusion on the health and well-being of First Nations people [3], who experience significantly higher rates of poverty, chronic disease, infectious disease, and mortality compared to the general Canadian population [4,5].

In Ontario, there are 133 First Nations communities, most of which have a reserve land base and have members living both on-reserve and off-reserve. These First Nations come together for collective decision-making, action, and advocacy through the Chiefs of Ontario (COO). In 2012, after exten-

*Corresponding Author:

Email Address: jenniferwalker@laurentian.ca (JD Walker)

sive negotiations and relationship building, COO entered into a data governance agreement (DGA) with the Institute for Clinical Evaluative Sciences (ICES), the principal independent steward of linked health data in Ontario. The DGA was designed to facilitate First Nations-engaged research and to ensure that Indigenous data sovereignty principles governing the use of First Nations data were grounded in the principles of OCAP®. The governance process for use of routinely collected health data with Indigenous identifiers at ICES has been published previously [6] and has been reproduced here (Figure 1). While this process is not the focus of the present paper, it has guided the way in which the data in question was linked at ICES. Through the COO-ICES agreement specifically, any use of data at ICES that directly or indirectly identifies First Nations people or communities is subject to First Nations governance processes. This ensures that all First Nations-specific analyses of data at ICES are undertaken according to First Nations collective priorities and apply Indigenous community-based research approaches.

ICES is an independent not-for-profit research institute encompassing a community of research, data and clinical experts, and a secure and accessible array of Ontario's health-related data. This consists of record-level, coded and linkable health data sets, including demographic and administrative records, registries, laboratory data and survey data. ICES links these files to create research-ready longitudinal de-identified person-level data-sets for a population of 13.5 million. The ICES data repository is used to conduct analyses of about many aspects of health care in Ontario, including: health conditions, health service experiences, health system performance, and patient outcomes.

One of the key barriers to Indigenous health data analysis in Ontario was the lack of comprehensive First Nations identifiers in the existing routinely-collected data at ICES, which do not consistently contain race or ethnicity fields. Using Ontario health administrative data alone, analyses of First Nations health outcomes were limited to geographic identification of First Nations people living on-reserve. To overcome this deficiency, ICES and COO jointly approached Indigenous and Northern Affairs Canada (INAC), a department of the Canadian federal government, to obtain a copy of the Indian Register (IR), which contains information on all First Nations persons in Canada who are recognized under the Indian Act. The ICES-COO DGA allowed for the acquisition and linkage of the IR, which served to essentially unlock the data at ICES for use by First Nations organizations and communities. The overall aim of this work was to create a First Nations cohort for health research in Ontario. This paper describes the technical process of linking the IR to the central data holdings at ICES to create this cohort within the context of Indigenous data sovereignty principles.

Methods

We linked the IR to the Ontario Registered Persons Database (RPDB) using both deterministic and probabilistic methods. This process identified Ontario-registered First Nations people in ICES data holdings, and allowed us to match them to their health and demographic records. This created the largest First Nations health research study cohort in Canada, which

is being used for disease surveillance and evaluation of health care.

Databases

The Government of Canada maintains a list of all registered First Nations people, living both on- and off-reserve, known as the IR. The eligibility criteria that a person must meet to be registered are set out in the *Indian Act* [7], which was first passed in 1876 and has been amended numerous times. The IR database contains demographic and administrative information including: individual identifiers for linkage (names, sex, date of birth), band affiliation, date of registration, record status (e.g. active, inactive, confirmed death), province of residence when registered, residence status (on reserve/off reserve) when registered, parent-child relationships, and marital status when registered.

The federal IR data were linked to the RPDB at ICES in 2014. The RPDB provides basic demographic information about anyone who has ever received an Ontario health card number (i.e. anyone who is, or who has ever been, eligible for health care in Ontario dating back to April 1, 1990). It contains basic demographic information including surname and first name, date of birth, sex, postal code, as well as a unique health card identifier, enabling linkage with other health utilization data. The Ontario Ministry of Health and Long-Term Care provides data updates to ICES monthly, and these are enriched by linkage to other ICES data holdings. Given Canada's universal health care coverage, the RPDB captures the majority of Ontario's 13.5 million residents. It is important to note that responsibility for health care for First Nations people is shared between the Federal and Provincial/Territorial governments in Canada, whereas Provincial/Territorial governments hold the responsibility for health care for the general population. As such, individuals living in far Northern and remote First Nation communities in Ontario with limited connection to the mainstream provincial health care system will not be well-represented in the Ontario health data. Nonetheless, those who have (or have ever received) an Ontario health card number will be in the ICES RPDB, regardless of where they live.

Data linkage

At ICES, record linkage of health records is commonly performed using the Ontario Health Insurance Program (OHIP) number. However, this information is not recorded in the federal IR. Consequently, we used the Automatch probabilistic record linkage program to link the IR records to the RPDB using deterministic and probabilistic approaches [8]. Deterministic matching methods used a combination of surname, given name, and date of birth, and require perfect agreement on these fields from both data files. Probabilistic data linkage is based on fields which may not be unique in both data files; in addition, there may be discrepancies in the information due to key entry errors or misspellings. In probabilistic data linkage, linkage weights are assigned to generate a theoretical likelihood that two records are a true match [9].

Linkage weights are based on two probabilities, the m probability, which is the conditional probability that a field agrees given the pair is a true match, and the u probability, the conditional probability that a field agrees given the pair is a true

Figure 1: Governance processes, reproduced with permission from Walker et al.¹**Panel: Governance processes for use of routinely collected health data with Indigenous identifiers at the Institute for Clinical Evaluative Sciences in Ontario, Canada.**

1. Access to, and use of data with Indigenous identifiers are approved by data governance committees organised and populated by the relevant Indigenous organisations.*
2. Linked datasets with Indigenous identifiers are not routinely available to researchers and analysts, who must make specific application, and seek approval from the relevant data governance committee before they can access them.[†]
3. Researchers are required to discuss their projects with Indigenous community representatives, who may collaborate in the planning conduct and reporting of the studies.
4. Researchers and staff at ICES participate in ongoing initiatives to orient them to Indigenous worldviews, research principles, and historical and social contexts.
5. Staff at ICES are working with representative organisations to build capacity among Indigenous organisations and communities to train Indigenous analysts and epidemiologists.
6. Study results are co-interpreted with the communities and their representatives, who have a lead role in deciding how the results will be communicated more widely.

*We describe the governance of First Nations data; similar arrangements have been established with the Métis Nation of Ontario for analysis of data that identify Métis individuals.

[†]Indigenous persons are routinely included in analyses of the whole province, or regions, but are not identified separately in the results. These analyses do not require clearance by the governance bodies. ICES= Institute for Clinical Evaluative Sciences, Ontario, Canada.

non-match. The true probabilities of matching cannot be estimated. Rather than representing a match probability, linkage weights are more accurately described as a match score. The higher the score, the greater the likelihood that the two records belong to the same individual, which itself depends on the assumption that the m and u probabilities are independent [10,11]. Two pre-determined thresholds were established, where matched pairs with linkage weights falling above the high threshold were considered automatic matches and pairs with weights falling below the low threshold were considered non-matches. Matched pairs with linkage weights between the two threshold weights were considered possible matches and were subject to manual review by a data covenantor in a room with restricted entry. In cases where multiple records from the IR data file were probabilistically matched to the same record from the RPDB, only the match with the highest linkage weight was considered.

We first stratified by sex to reduce the total number of comparisons. One pass using deterministic matching methods was followed by six subsequent probabilistic passes. Surnames were standardized using the New York State Identification and Intelligence System (NYSIIS) phonetic conversion [12]. With data files of this size, it would not be feasible to scan every record for matched pairs. A technique called *blocking* partitioned each file into mutually exclusive and exhaustive subsets, and we looked for matches within each subset. This method greatly reduces the number of possible pairs that are scanned

for matches. If a match could not be confirmed after the first pass, the process continued to look for matches by utilizing different probabilistic blocking schemes. The description of the block used for each pass are shown in Table 1. Successfully linked records from the IR file were assigned unique ICES key numbers (IKN), which are derived from the OHIP numbers. In parallel to the data linkage, we conducted a file unduplication of the IR to estimate the proportion of duplicate registrants within the file.

There is no established population of First Nations people living in Ontario with which we could establish a linkage rate. As such, we applied several approaches to determining a denominator that would represent the total population of First Nations we would hope to link to the RPDB. One approach was to display the linkage rate for only those records in the national IR database with an Ontario band number. This limited the assessment of the linkage to those who were affiliated with a First Nation community in Ontario. However, individuals can be affiliated with a community in Ontario but never live or receive health services in Ontario. The second approach was to assess the linkage in only those records in the IR that had an Ontario province code (which is the individual's province of residence at the time of registration in the IR database). The other two comparisons that we made were to external sources of population data for First Nations in Ontario: the 2006 Census and the First Nations profiles from the Government of Canada's website. Finally, we looked at the

¹Reprinted from The Lancet, Volume 390, Walker J, Lovett R, Kukutai T, Jones C, Henry D., Indigenous health data and the path to healing., Page 2022, Copyright (2017), with permission from Elsevier.

linkage rates over time for those who were registered with a First Nations community in Ontario.

Results

The results of the linkage using both deterministic and probabilistic matching methods are shown in Table 1, by sex and linkage pass number. There were 1,027,973 individuals in the national copy of the IR that we received from INAC. In total, 201,678 individuals matched to Ontario health records by way of the RPDB, of which 98,562 were female and 103,116 were male. The IR unduplication procedure involved five passes; one pass using deterministic matching methods, followed by four probabilistic passes. Based on the matching methods employed, 0.32% of the IR records are possible duplicates.

Table 2 displays the linkage results for those First Nations people who lived in Ontario at the time when they registered with the IR ($n=213,233$) and shows that 176,266 of these individuals were successfully linked to the RPDB. Table 2 also displays the linkage results for those individuals who are registered with a First Nations community in Ontario ($n=193,444$) and shows that 149,728 (77.4%) of these records linked.

Of those First Nations individuals linked to the RPDB ($n=201,678$), 87.4% lived in Ontario when they first registered with IR (ON province code); 74.2% were affiliated with an Ontario First Nation Community (ON band number); 90.2% had an Ontario province code or an Ontario band number; and 9.8% had neither.

Table 3 shows the number of males and females in the linked IR file, relative to other data sources including the “North American Indian” Identity question and self-identification of “Registered Indian” status on the 2006 Census [13], as well as the First Nations Profiles from the INAC website [14].

The proportion of registered First Nations people linking to the RPDB improved across time (year of registration), from 62.8% in the 1960s to 94.5% in 2012 (Figure 1). Of the 43,907 records missing the year of registration information, 71.7% linked to the RPDB.

Discussion

The collaboration between COO and ICES enabled linkage of the IR to the Ontario RPDB, creating the largest First Nations health research study cohort in Canada. This was accomplished by building a mutually respectful partnership that is strengthened and supported by a DGA that establishes First Nations data governance and OCAP® principles of First Nations ownership, control, access and possession of First Nations data. With this core linkage between IR and RPDB, the vast array of health system, mortality and demographic data at ICES is available for high quality, First Nations-directed research. This will be an invaluable tool for First Nations as they work to build strong and healthy populations.

This linkage was technically challenging on several fronts. It is customary for record linkage studies to provide a “linkage rate”, which may be used to describe the proportion of the records in a new database (i.e. the IR) that are successfully linked to the other database (i.e. the RPDB). However, the

ability to provide a linkage rate for this study is substantially limited by the fact that the IR is a *federal* database containing records of all First Nations people in Canada, while the RPDB is a *provincial* database. Thus, the main challenge lies with defining who is an Ontario resident.

An obvious approach is to look at those with an Ontario province code recorded in the IR; yet, this simply means that an individual was living in Ontario at the time of their registration. For many people, registration occurs early in life, often around the time of birth. Accordingly, people with an Ontario province code may no longer live in Ontario. This provides some explanation for the discrepancy between the 213,233 people with an Ontario province code in the entire IR file, and the 176,266 people with an Ontario province code linked to the RPDB.

Another approach is to look at those with an Ontario band number recorded in the IR, representing the First Nations community with which each person is affiliated. An individual may have clear and strong familial, social, and spiritual ties to their community, but that does not necessarily mean that they reside – or have ever resided – in Ontario. As such, the Ontario band number also has its shortcomings as an indicator of success of the IR linkage; 77.4% of the 193,444 people with an Ontario band number linked to the RPDB. Table 2 also presents the proportion of those with Ontario band number or province code as a proportion of the total number people linked to the RPDB ($n=201,678$), which results in the highest proportion of those who successfully linked (90.2%). Due to the limitations of each of these approaches, we cannot present a true linkage rate.

The challenges of using an Ontario province code or band number to define denominators for calculating linkage rates are complicated by the fact that approximately 10% ($n=19,763$) of individuals in the linked IR file have no recorded information on either variable. This is plausible given that a person may (i) have lived outside of Ontario at the time of registration (non-Ontario province code), (ii) belong to a First Nations community outside of Ontario (non-Ontario band number), and (iii) have subsequently moved to Ontario and registered with the Ontario Health Insurance Plan.

When attempting to define who is an Ontario resident, it is important to consider that where First Nations people live may not be where they access health services. The “reserve lands” located across Ontario are a product of the Indian Act and the federal government’s system of segregation [15]. Many of these First Nations communities lie close to provincial borders, with a particularly populated area being Northwestern Ontario, bordering Manitoba. It is not uncommon for First Nations people residing in Ontario to access health services in Manitoba and vice-versa. The need to travel outside one’s province for appropriate and timely health care points to a health care access issue also seen at Ontario-Quebec and Ontario-U.S. borders. More work to link data from multiple provinces is warranted to better capture the health service use and health outcomes of First Nations in Ontario and in other provinces.

The 2006 Census was used for comparison due to a change to the voluntary National Household Survey in 2011. The 2006 Census had distinct questions asking respondents if they self-identify as “North American Indian” ($n=158,400$) and as a “Registered Indian” ($n=123,595$), with results for both indi-

Table 1: Results of IR Linkage to the RPDB, using Deterministic and Probabilistic Matching Methods, Number (percent of total linked records).

| Linkage Pass Number | Description of Block | Females | Males | Total |
|---------------------------------------|---|-------------------|-------------------|--------------------|
| Deterministic Matching Methods | | | | |
| 1 | Surname (Main) + Given Name 1 + Date of Birth | 85,931 (87.18) | 89,781 (87.07) | 175,712 (87.13) |
| Probabilistic Matching Methods | | | | |
| 2 | Surname Initial + Given Name 1 Initials (1st-3 Chars) + Date of Birth | 6,970 (7.07) | 7,088 (6.87) | 14,058 (6.97) |
| 3 | DOB | 4,694 (4.76) | 4,427 (4.29) | 9,121 (4.52) |
| 4 | Surname Initial + Given Name 1 Initials (1st-3 Chars) + Birth Year | 548 (0.56) | 1,218 (1.18) | 1,766 (0.88) |
| 5 | Given Name 1 Initials (1st-3 Chars) + Birth Month + Birth Day | 263 (0.27) | 436 (0.42) | 699 (0.35) |
| 6 | NYSIIS + Birth Year | 139 (0.14) | 131 (0.13) | 270 (0.13) |
| 7 | Surname Initial + Birth Month + Birth Day | 17 (0.02) | 35 (0.03) | 52 (0.03) |
| Total # of Matches | | 98,562 | 103,116 | 201,678 |
| Total # of IR Records | | | | 1,027,973 |

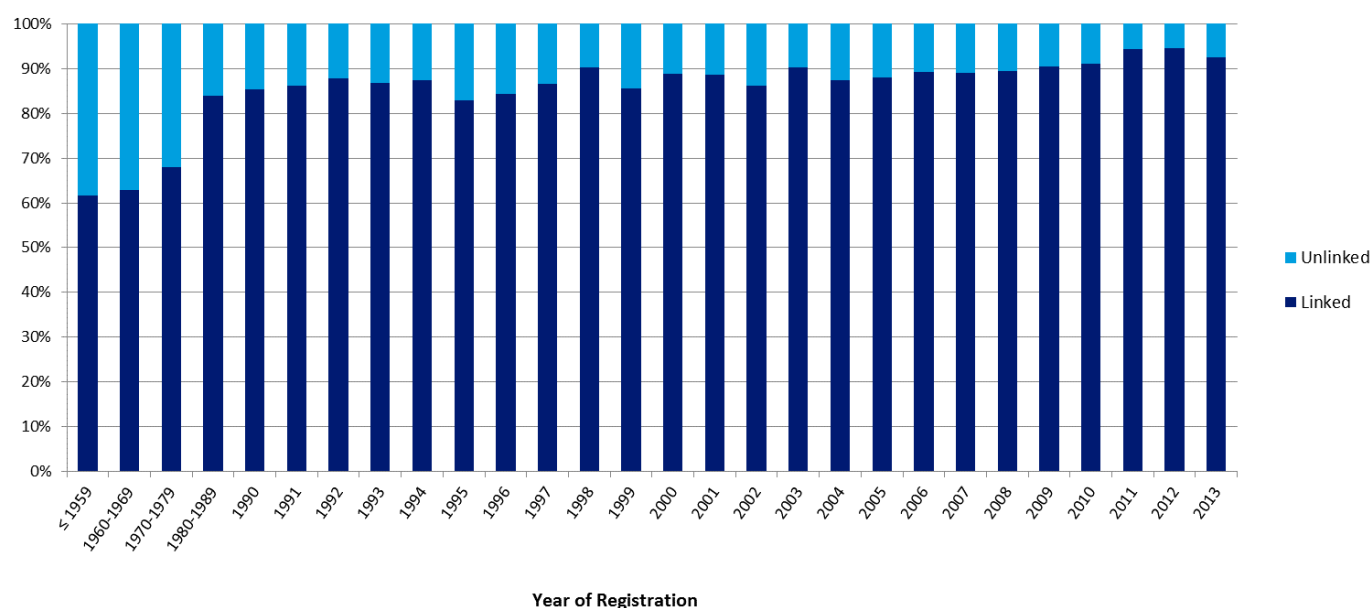
Table 2: Linkage results by Ontario band number and Ontario province code

| | ON Province Code | ON Band Number | ON Province Code or ON Band Number | No ON Province Code nor ON Band Number | Total |
|---|-----------------------------|---------------------------|---------------------------------------|---|-----------|
| Linked with RPDB (row %) (column %) | 176,266 (87.4) (82.7) | 149,728 (74) (77.4) | 181,915 (90) (77.9) | 19,763 (10) (2.5) | 201,678 |
| Total in unlinked IR file | 213,233 | 193,444 | 233,611 | 794,362 | 1,027,973 |

Table 3: Comparison of sources intending to capture First Nations population counts, by sex

| | IR-linked file | 2006 Census: "North American Indian" Identity [9] | 2006 Census: Self-identified "Registered Indian" [9] | Government of Canada Website [10] |
|----------------|----------------|---|--|-----------------------------------|
| Females | 98,562 | 82,445 | 64,815 | 105,453 |
| Males | 103,116 | 75,955 | 58,780 | 99,601 |
| Total | 201,678 | 158,400 | 123,595 | 205,054 |

Figure 2: Proportion of those with an Ontario province code linked and unlinked, by year of registration in the IR



cating more females than males. A clear limitation of using self-identifiers is the reluctance of many First Nations people to self-identify. Table 3 reveals that the self-identifiers have lower counts than the IR-linked file. It is important to note that in 2006, 22 First Nations communities ("Indian reserves and Indian settlements") were incompletely enumerated by the census or declined to participate and the populations of these 22 communities were not included in the census counts [13]. This incomplete enumeration also partially accounts for the difference between the Census and INAC counts. The limited utility of the census estimates for Indigenous populations demonstrated here and elsewhere [16] underscores the strong need for Indigenous-led collection and management of population data.

An alternative source of data on denominators are the First Nations population counts retrieved from the Government of Canada website in 2017. They include registered First Nations, both on- and off-reserve for Ontario bands. In contrast, the IR-linked file may include individuals from any First Nations community (i.e. any band number) in Canada, so long as they have an Ontario health card. Another distinction between the two data sources is that the Government of Canada values ostensibly include those registered up to 2017, while the IR was linked in 2014, resulting in incomplete enumeration beyond 2013.

Finally, Table 3 highlights that while other sources show more First Nations females than males, the IR-linked file shows the opposite. In most probabilistic data linkages, there is a lower rate of female matches compared to male matches—a phenomenon mainly attributable to surname changes, which are much more common in women than in men. An additional contributing factor is likely gender discrimination in the IR; until revisions to the Indian Act in 1985, women who married non-status men lost their status as registered First Nations people.

Despite lacking an ideal denominator of First Nations in

Ontario, and thus not having a standard linkage rate, we have been able to compare results of the IR linkage using various approaches, as well as draw comparisons with other data sources. By characterizing the linked and unlinked populations, we can identify issues of bias in subsequent analyses. Ultimately, the more work that can be done to understand the limitations of the data, the better equipped we will be to utilize existing data sources to generate a comprehensive and representative picture of First Nations health.

Limitations

The IR and RPDB data linkage aims to encompass the First Nations population in Ontario; however, the IR data does not identify non-status First Nations people, who are not registered under the Indian Act [7]. There are several reasons why a First Nations person may not be registered that relate to past and current criteria for registration established by the Canadian government. For example, a First Nations person may not be registered if they had an ancestor who was not physically present at the time and place of treaty signing and registration or because their community did not sign a treaty with the Crown. In addition, some people who were initially registered were subsequently removed from the IR under the legal authority of the Indian Act through compulsory enfranchisement or loss of Indian Status [17, 18]. Examples of compulsory enfranchisement include status women who married non-status men prior to 1985, and First Nations people who attended university, joined the army, or chose to vote in federal elections prior to 1960 [3,17,18]. As such, there are members of First Nations communities and families who cannot or chose not to be registered under the Indian Act.

There are also limitations of the RPDB that must be considered. First, it does not capture First Nations people who have had no contact with the Ontario health system. This would include individuals who do not have an Ontario health

card, and may also include those who primarily receive health care through a federal nursing station in a remote community or who primarily seek health care in a different province. Second, because homelessness is a concern for First Nations people, particularly those living in cities [19], more work needs to be done to understand the under-representation of homeless individuals in the linked data.

Third, the RPDB relies predominantly on the address associated with the health card to determine the place of residence. Many First Nations people frequently move between their home communities and urban settings [20] and this may result in a misclassification of people living on- or-off reserve and in specific geographic areas.

Conclusion

In summary, the Chiefs of Ontario and ICES have worked together to create the governance and technical infrastructure for extensive and appropriate use of First Nations-identified health systems data in Ontario. This has resulted in the creation of the largest First Nations health research study cohort in Canada with a high level of linkage accuracy, ranging from 77.4% to 82.7%. First Nations-driven research priorities are now being addressed using the linked data. These priorities span life stages (aging, child health), health care experiences (trajectories of care for diabetes), and mortality (preventable and premature deaths). The IR and RPDB linkage, though limited in some ways, is undoubtedly critical for our collective ability to answer questions from First Nations communities about the health of their people as they work towards higher levels of wellbeing and healing.

Acknowledgements

This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred.

Statement on conflicts of interest

The authors have no real or perceived conflicts of interest.

Abbreviations

| | |
|--------|---|
| AANDC | Aboriginal Affairs and Northern Development Canada |
| COO | Chiefs of Ontario |
| ICES | Institute for Clinical Evaluative Sciences |
| IKN | ICES Key Number |
| INAC | Indigenous and Northern Affairs Canada |
| IR | Indian Register |
| NYSIIS | New York State Identification and Intelligence System |
| OCAP® | Ownership, Control, Access, Possession |
| ON | Ontario |
| RPDB | Registered Persons Database |

References

1. Kukutai T, Taylor J. Data sovereignty for Indigenous peoples: current practice and future needs. In: Kukutai T, Taylor J, eds. *Indigenous data sovereignty: towards an agenda*. Research monograph no 38. ANU Press. 2016 Nov;38:1–18. Available from: <https://doi.org/10.22459/CAEPR38.11.2016>
2. The First Nations Information Governance Centre. Ownership, Control, Access and Possession (OCAP®): The Path to First Nations Information Governance. FNIGC. 2014 May. Available from: http://fnigc.ca/sites/default/files/docs/ocap_path_to_fn_information_governance_en_final.pdf
3. Truth and Reconciliation Commission of Canada. Honouring the truth, reconciling for the future: summary of the final report of the Truth and Reconciliation Commission of Canada. 2015. Available from: http://nctr.ca/assets/reports/Final%20Reports/Executive_Summary_English_Web.pdf
4. Tjepkema M, Wilkins R, Pennock J, Goedhuis N. Potential years of life lost at ages 25 to 74 among Status Indians, 1991 to 2001. *Health reports*. 2011 Mar 1;22(1):25. Available from: http://www.statcan.gc.ca/access_acces/alternative_alternatif.action?l=eng&loc=11409-eng.pdf
5. Mosby I, Galloway T. “Hunger was never absent”: How residential school diets shaped current patterns of diabetes among Indigenous peoples in Canada. *CMAJ*. 2017 August 14;189:E1043-5. Available from: <http://www.cmaj.ca/content/189/32/E1043https://doi.org/10.1503/cmaj.170448>
6. Walker J, Lovett R, Kukutai T, Jones C, Henry D. Indigenous health data and the path to healing. *Lancet*. 2017 Nov 4;390(10107):2022-23. [https://doi.org/10.1016/S0140-6736\(17\)32755-1](https://doi.org/10.1016/S0140-6736(17)32755-1)
7. Indian Act. In *Revised Statutes of Canada*. Queens Printer for Canada. 1985, c. 1-5. Available from: <http://laws-lois.justice.gc.ca/eng/acts/I-5/>
8. Jaro M. Automatch: Generalized Record Linkage System, Version 4.2: User Manual. Kennebunk, Maine: Matchware Technologies, Inc. Silver Spring: Matchware Technologies Inc.; 1994.
9. Jaro M. Probabilistic linkage of large public health data files. *Stat Med*. 1995 Mar 15-Apr 15;14(5-7):491-8. <https://doi.org/10.1002/sim.4780140510>
10. Doidge JC, Harron K. Demystifying probabilistic linkage: Common myths and misconceptions. *International Journal for Population Data Science*. 2018 3:1. <https://doi.org/10.23889/ijpds.v3i1.410>
11. Winkler WE. Overview of record linkage and current research directions. Washington, D.C.: U.S. Census Bureau; 2006. Available from: <https://www.census.gov/srd/papers/pdf/rrs2006-02.pdf>

12. Taft R. Name search techniques: New York state identification and intelligence system. Special Rep; 1970. p. 1.
13. Statistics Canada. 2007. Ontario (Code35) [Table]. Aboriginal Population Profile. 2006 Census. Statistics Canada Catalogue no. 92-594-XWE. 2007 (Released 2008 Jan 15). Available from: <http://www12.statcan.ca/census-recensement/2006/dp-pd/prof/92-594/index.cfm?Lang=E>
14. Indigenous and Northern Affairs Canada [Internet]. First Nations Profiles. 2017. Available from: <http://fnp-ppn.aandc-aadnc.gc.ca/fnp/Main/Index.aspx?lang=eng>
15. Allan B, Smylie, J. First Peoples, second class treatment: The role of racism in the health and well-being of Indigenous peoples in Canada. The Wellesley Institute. 2015. Available from: <http://www.wellesleyinstitute.com/wp-content/uploads/2015/02/Summary-First-Peoples-Second-Class-Treatment-Final.pdf>
16. Smylie J, Firestone M. Back to the basics: Identifying and addressing underlying challenges in achieving high quality and relevant health statistics for indigenous populations in Canada. Stat J IAOS. 2015;31(1):67-87. Available from: <https://doi.org/10.3233/SJI-150864>
17. Canada. Royal Commission on Aboriginal Peoples. People to people, nation to nation: Highlights from the report of the Royal Commission on Aboriginal Peoples. Minister of Supply and Services Canada. 1996. Available from: <http://www.aadnc-aandc.gc.ca/eng/1100100014597/1100100014637>
18. Elections Canada. A History of the Vote in Canada. Chapter 3: Modernization, 1920–1981. 2017, c.3. Available from: <http://www.elections.ca/content.aspx?section=res&dir=his&document=chap3&lang=e>
19. Smylie J, Firestone M, Cochran L, Prince C, Maracle S, Morley M, Mayo S, Spiller T, McPherson B. Our Health Counts. Urban Aboriginal Health Data Research Project. Community Report: First Nations Adults and Children. 2011 Apr. Available from: <http://www.stmichaelshospital.com/pdf/crich/our-health-counts-report.pdf>
20. Chiefs of Ontario. First Nations Regional Health Survey (RHS) Ontario Region Phase 2 (2008/10): People's Report. 2014 Dec. Available from: <http://health.chiefs-of-ontario.org/sites/default/files/files/RHS%20Phase%20-%20-%20Peoples%20Report.pdf>

